Taylor Larrechea
Dr. Gustafson
MATH 366
CP 19.2

1) To solve an equation when there is no formula for the exact solution available, we can use an approximation method, such as an iteration method. This is a method in which we start from an initial guess $x_0$ (which may be poor) and compute step by step (in general better and better) approximations $x_1, x_2, \dots$ of an unknown solution.

2) By some algebraic steps we transform (1) into the form

(2) $$x = g(x)$$

Then we choose an $x_0$ and compute $x_1 = g(x_0)$, $x_2 = g(x_1)$, and in general

(3) $$x_{n+1} = g(x_n) \qquad (n = 0, 1, \cdots)$$

A solution of (2) is called a fixed point of $g$, motivating the name of the method. This is a solution of (1), since from $x = g(x)$ we can return to the original form $f(x) = 0$. From (1) we may get several different forms of (2). The behavior of corresponding iterative sequences $x_0, x_1, \dots$ may differ, in particular, with respect to their speed of convergence. Indeed, some of them may not converge at all. Let us illustrate these facts with a simple example.

3) (a) $\dfrac{-b \pm \sqrt{b^2 - 4ac}}{2a}$ : $a = 1, b = -3, c = 1$ : $x_1 = \dfrac{3 + \sqrt{9 - 4(1)(1)}}{2(1)}$, $x_2 = \dfrac{3 - \sqrt{9 - 4(1)(1)}}{2(1)}$

$$x_1 = \frac{3 + \sqrt{5}}{2} = 1.5 + \sqrt{1.25}, \quad x_2 = \frac{3 - \sqrt{5}}{2} = 1.5 - \sqrt{1.25}$$

$$\boxed{x_1 = 1.5 + \sqrt{1.25},\ x_2 = 1.5 - \sqrt{1.25}}$$

(b) $f(x) = x^2 - 3x + 1 = 0$

$x^2 - 3x + 1 = \ \to\ x^2 + 1 = 3x \ \to\ \frac{1}{3}(x^2 + 1) = x$ Thus: $x_{n+1} = \frac{1}{3}(x_n^2 + 1)$

$x^2 - 3x + 1 = \ \to\ x^2 = 3x - 1 \ \to\ x = 3 - 1/x$ Thus: $x_{n+1} = 3 - 1/x_n$

(c) Our figures show the following. In the lower part of Fig. 426a the slope of $g_1(x)$ is less than the slope of $y = x$, which is 1, thus $|g_1'(x)| < 1$, and we seem to have convergence. In the upper part, $g_1(x)$ is steeper $(g_1'(x) > 1)$ and we have divergence. In Fig. 426b the slope of $g_2(x)$ is less near the intersection point ($x = 2.618$, fixed point of $g_2$, solution of $f(x) = 0$), and both sequences seem to converge. From all this we conclude that convergence seems to depend on the fact that, in a neighborhood of a solution, the curve of $g(x)$ is less steep than the straight line $y = x$, and we shall now see that this condition $|g'(x)| < 1$ (= slope of $y = x$) is sufficient for convergence.

4) Theorem 2: Convergence of Fixed-Point Iteration

Let $x=s$ be a solution of $x=g(x)$ and suppose that $g$ has a continuous derivative in some interval $J$ containing $s$. Then, if $|g'(x)| \leqq K < 1$ in $J$, the iteration process defined by (3) converges for any $x_0$ in $J$. The limit of the sequence $\{x_n\}$ is $s$.

5) Algorithm Newton ($f, f', x_0, \epsilon, N$)

This algorithm computes a solution of $f(x)=0$ given an initial approximation $x_0$ (starting value of the iteration.) Here the function $f(x)$ is continuous and has a continuous derivative $f'(x)$.

INPUT: $f, f'$, initial approximation $x_0$, tolerance $\epsilon > 0$, maximum number of iterations $N$.
OUTPUT: Approximate solution $x_n$ ($n \leqq N$) or message of failure.

For $n = 0, 1, 2, \ldots, N-1$ do:

1 Compute $f'(x_n)$

2 If $f'(x_n) = 0$ then OUTPUT "Failure." Stop.
[Procedure completed unsuccessfully]

3 Else compute

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \tag{5}$$

4 If $|x_{n+1} - x_n| \leqq \epsilon |x_{n+1}|$ then OUTPUT $x_{n+1}$. Stop.
[Procedure completed successfully]

5 OUTPUT "Failure". Stop.
[Procedure completed unsuccessfully after $N$ iterations]

END

END Newton

6) The inequality in line 4 is a termination criterion. If the sequence of the $x_n$ converges and the criterion holds, we have reached the desired accuracy and stop. Note that this is just a form of the relative error test. It ensures that the result has the desired number of significant digits. If $|x_{n+1}| = 0$, the condition is satisfied if and only if $x_{n+1} = x_n = 0$, otherwise $|x_{n+1} - x_n|$ must be sufficiently small. The factor $|x_{n+1}|$ is needed in the case of zeros of very small (or very large) absolute value because of the high density (or of the scarcity) of machine numbers for those $x$.

7) Line 5 gives another termination criterion and is needed because Newton's method may diverge or, due to a poor choice of $x_0$, may not reach the desired accuracy by a reasonable number of iterations. Then we may try another $x_0$. If $f(x)=0$ has more than one solution, different choices of $x_0$ may produce different solutions. Also, an iterative sequence may sometimes converge to a solution different from the expected one.

8) Example 5 Newton's Method Applied to an Algebraic Equation

Apply Newton's method to the equation $f(x) = x^3 + x - 1 = 0$.
Solution. From (5) we have

$$x_{n+1} = x_n - \frac{x_n^3 + x_n - 1}{3x_n^2 + 1} = \frac{2x_n^3 + 1}{3x_n^2 + 1}.$$

Starting from $x_0=1$, we obtain

$$x_1 = 0.750000, \quad x_2 = 0.686047, \quad x_3 = 0.682340, \quad x_4 = 0.682328$$

where $x_4$ has the error $-1\cdot 10^{-6}$. A comparison with Example 2 shows that the present convergence is much more rapid. This may motivate the concept of the order of an iteration process, to be discussed next.

## 9) Order of an Iteration Method. Speed of Convergence.

The quality of an iteration method may be characterized by the speed of convergence, as follows.

Let $x_{n+1}=g(x_n)$ define an iteration method, and let $x_n$ approximate a solution $s$ of $x=g(x)$. Then $x_n = s-\epsilon_n$, where $\epsilon_n$ is the error of $x_n$. Suppose that $g$ is differentiable a number of times, so that the Taylor formula gives

(6) $$x_{n+1}=g(x_n)=g(s)+g'(s)(x_n-s)+\frac{1}{2}g''(s)(x_n-s)^2+\ldots\ldots = g(s)-g'(s)\epsilon_n+\frac{1}{2}g''(s)\epsilon_n^2+\ldots$$

The exponent of $\epsilon_n$ in the first nonvanishing term after $g(s)$ is called the order of the iteration process defined by $g$. The order measures the speed of convergence.

To see this, subtract $g(s)=s$ on both sides of (6). The on the left you get $x_{n+1}-s=-\epsilon_{n+1}$, where $\epsilon_{n+1}$ is the error of $x_{n+1}$. And on the right the remaining expression equals approximately its first nonzero term because $|\epsilon_n|$ is small in the case of converge. Thus

(7)
(a) $\epsilon_{n+1} \approx +g'(s)\epsilon_n$ in the case of first order,

(b) $\epsilon_{n+1} \approx -\frac{1}{2}g''(s)\epsilon_n^2$ in the case of second order, etc.

Thus if $\epsilon_n = 10^{-k}$ in some step, then for second order, $\epsilon_{n+1}=c\cdot(10^{-k})^2 = c\cdot 10^{-2k}$, so that the number of significant digits is about doubled in each step.

## 10) Theorem 2: Second-order Convergence of Newton's Method

If $f(x)$ is three times differentiable and $f'$ and $f''$ are not zero at a solution $s$ of $f(x)=0$, then for $x_0$ sufficiently close to $s$, Newton's method is of second order.

Comments: For Newton's method, (7b) becomes, by (8*),

(9) $$\epsilon_{n+1} \approx -\frac{f''(s)}{2f'(s)}\epsilon_n^2$$

For the rapid convergence of the method indicated in Theorem 2 it is important that $s$ be a simple zero of $f(x)$ (thus $f'(s)\neq 0$) and that $x_0$ be close to $s$, because in Taylor's formula we took only the linear term [see (5*)], assuming the quadriatic term to be negligibly small. (with a bad $x_0$ the method may even diverge!)

## 11) Difficulties in Newton's Method:

Difficulties may arise if $|f'(x)|$ is very small near a solution $s$ of $f(x)=0$. In this case of $f(x)=0$, we call $f(x)$ ill-conditioned.

## 12) Secant Method for Solving $f(x)=0$

Newton's method is very powerful but has the disadvantage that the derivative $f'$ may sometimes be a far more difficult expression than $f$ itself and its evaluation therefore computationally expensive. This situation suggests the idea of replacing the derivative with the difference quotient

$$f'(x_n) \approx \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}}$$

Then instead of (5) we have the formula of the popular Secant method

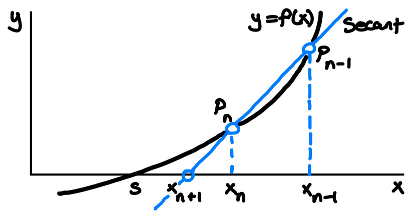


Fig. 429. Secant method

(10) $$x_{n+1} = x_n - f(x_n)\frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})}$$

Geometrically, we intersect the x-axis at $x_{n+1}$ with the secant of $f(x)$ passing through $P_{n-1}$ and $P_n$ in Fig. 429. We need two starting values $x_0$ and $x_1$. Evaluation of derivatives is now avoided. It can be shown that convergence is superlinear (that is, more rapid than linear, $|\epsilon_{n+1}| \approx \text{const} \cdot |\epsilon_n|^{1.62}$; see [E5] in App. 1), almost quadratic like Newton's method. The algorithm is similar to that of Newton's method, as the student may show.

CAUTION! It is not good to write (10) as

$$x_{n+1} = \frac{x_{n-1} f(x_n) - x_n f(x_{n-1})}{f(x_n) - f(x_{n-1})}$$

because this may lead to loss of significant digits if $x_n$ and $x_{n-1}$ are about equal. (Can you see this from the formula?)